Volume 13 Number 5 1985 Nucleic Acids Research

Structural organization of the P_{25} gene of *Bombyx mori* and comparative analysis of its 5' flanking DNA with that of the fibroin gene

Pierre Couble, Martine Chevillard, Agnès Moine, Patrick Ravel-Chapuis and Jean-Claude Prudhomme

Département de Biologie Générale et Appliquée, Laboratoire Associé au CNRS no 92, Université Claude Bernard, Lyon I, 43, Boulevard du 11 Novembre 1918, Villeurbanne Cedex, France

Received 29 August 1984; Revised and Accepted 4 February 1985

ABSTRACT

We have cloned a large portion of the P_{25} gene of Bombyx mori encoding the 25,000 dalton polypeptide which associates with fibroin to constitute the major silk protein. Its structure has been investigated by restriction mapping R-loop analysis, S1 nuclease protection experiments and nucleotide sequencing of the region spanning the 5' end of the gene and its flanking DNA. This has permitted a comparative sequence analysis of the DNA from the P_{25} and fibroin genes. The genes demonstrate no relatedness in their coding regions but they exhibit large blocks of sequence homology in their 5' flanking regions. In particular, the DNA upstream of the P_{25} gene possesses a sequence very similar to a region of fibroin 5' flanking DNA that is known to possess transcription modulation signals. The functional significance of these homologous regions is discussed with regard to the highly coordinated expression of these two genes.

INTRODUCTION

P_{25} mRNA of Bombyx mori codes for a 25 kd silk protein which binds through disulfide bonds to fibroin, the major 370 kd component of silk (1,2). Quantitative measurements of P_{25} mRNA during development have shown that its production by posterior silkgland cells is restricted to specific stages of larval life (1,3). The mRNA is destroyed at molting; it then accumulates during the subsequent intermolt through intense transcription. Interestingly, the variations in the cellular content of P_{25} mRNA closely follow those of fibroin mRNA. Thus, both kinds of mRNAs accumulate in equi-molar amounts at the fifth intermolt, up to the onset of spinning (1). At this stage, these two mRNAs represent 80-85% of the mass of total mRNA. Yet P_{25} and fibroin mRNA are unrelated structurally : they differ in size, 1100 and 16000 nucleotides respectively (1,4), and in sequence since P_{25} cDNA does not hybridize fibroin mRNA and vice versa (3), and their protein products have very different amino acid composition (5).

Our interest in this system is to unravel the mode of expression of P_{25} and fibroin genes which function in a highly coordinated way despite the lack of any obvious evolutionary relationship between their coding regions. The

equimolarity observed throughout development between P_{25} and fibroin mRNA likely results from adjusted transcription rates. In support of this view, direct measurements of *in vivo* fibroin mRNA synthesis have shown that fibroin gene expression is modulated at the transcriptional level (6). Accordingly, we hypothesized that P_{25} and fibroin genes would share some common regulatory sequences in their 5' flanking regions, where control elements have been found in a variety of RNA polymerase II transcribed genes (7-11). The structural organization of the fibroin gene has been described by Y. Suzuki and his colleagues (12,13) and certain flanking sequences have been identified that have roles in promoting or modulating *in vitro* transcription (14,15). In this paper, we report the cloning and sequencing of a large portion of the P_{25} gene and we compare its flanking DNA with that of fibroin gene. The two genes share blocks of homologous sequences in their 5' flanking regions. Interestingly, the most striking homologies are found to correspond precisely to the region of the fibroin gene previously shown to contain strong signals for transcription modulation.

MATERIALS AND METHODS

Silkworms. Hybrids of *B. mori* European strains 200 and 300 were reared in the laboratory as described previously (3).

Enzymes. AMV reverse transcriptase was a gift from Dr J. Beard. Restriction enzymes were from BRL and Boerhinger. T4 DNA ligase, the Klenow fragment of DNA polymerase, T4 polynucleotide kinase, and calf intestinal phosphatase were purchased from Boerhinger. S1 nuclease was obtained from Miles and RNAsin from Biotech.

Nucleic acid isolation and hybridization. Procedures for the isolation of total RNA and poly A^+ RNA, and sucrose gradient separation of fibroin mRNA and non-fibroin mRNA were as described (3). P_{25} mRNA was purified by agarose gel electrophoresis (1). Liquid hybridizations were carried out as reported, using SI nuclease assays for quantitative detection of hybrids (3). DNA and RNA blots were hybridized for 18 to 48 hours at 42°C in 50% formamide, 5 x SSC, I x Denhardt's solution, 100 µg/ml denatured herring sperm DNA with nick-translated DNA or with cDNA. After hybridization, filters were washed stepwise with 2 x SSC, 0.2 x SSC and 0.1 x SSC containing 0.1% SDS at 65°C with gentle agitation. Autoradiography used Kodak XR5 films.

Recombinant DNA. The *B. mori* genomic library of strain 103, consisting of Eco RI fragments inserted into λCharon 4 (16), was kindly provided by T. Eickbush. Screening and isolation of recombinant DNA were carried out according to the plaque hybridization method (17). Hybridization probes were ^{32}P-labeled cDNA

from total mRNA or from agarose gel purified P_{25} mRNA. Recombinant phages were carried through 4 cycles of single plaque purification prior to amplification. Genomic restricted fragments of particular interest were subcloned in either the plasmid pBR322 ($pBmP_{25}21$) or the plasmid pACYC184 ($pBmP_{25}11$).

R-loop analysis. R-loops were formed (18) between P_{25} mRNA and recombinant DNA, the DNA being spread for electron microscopy using the formamide technique (19). 15 molecules were analysed for estimates of intron and exon lengths.

cDNA elongation. The conditions for cDNA synthesis have been described already (3). Full length cDNA was prepared in a 45 min reaction at 37°C whereas cDNA representing the 3' end of mRNA was obtained in a 4 min reaction at the same temperature. For cDNA sequencing, short restriction fragments were used as sources of end-labeled primer. Fragments were labeled at the 5' termini using γ-^{32}P-ATP and T4 polynucleotide kinase and were strand separated by polyacrylamide gel electrophoresis. 10-30 pmoles of single stranded primer were then combined with 10 µg of total mRNA from posterior silkgland and annealed in 0.4M NaCl, 1mM EDTA, 10 mM Tris-HCl (pH 7.9), 0.1% SDS for 2h at 65°C. The hybridization mixture was then diluted and adjusted for a 30 min reverse transcriptase reaction at 37°C. The resulting end-labeled cDNA was extracted by standard procedures for sequencing.

Sequence determination. Nucleotide sequences were determined by the methods of Maxam and Gilbert (20) and Sanger et al (21). For the latter, fragments of interest were subcloned in the phages M13 mp8, mp9, mp10 or mp11. The resulting single-stranded templates were labeled with α-^{32}P-dCTP or α-^{35}S-dATP (Amersham). The sequencing strategy is summarized in Figure 2. Sequence homology searches utilized the extensive set of programmes in the ACNUC bank (Institut d'Evolution Moléculaire, Lyon). The programme of Papanicolaou et al (22) was used to assess the stability of DNA secondary structures.

RESULTS

Cloning and identification of P_{25} gene recombinants

Three types of cDNA served as probes for the differential selection of phage carrying genomic P_{25} DNA fragments : a) cDNA from non-fibroin mRNA of posterior silkgland (PSG), which contains 25-30% of P_{25} cDNA and a very minor fraction of contaminating fibroin cDNA (3); b) cDNA transcribed from agarose gel-purified P_{25} mRNA containing 65% of the relevant sequence and no detectable fibroin cDNA contaminant (1); c) cDNA from total mRNA of the middle silkgland (MSG) whose cells do not express P_{25} or fibroin genes (1,23). The library screened was that constructed by T. Eickbush, comprising a partial

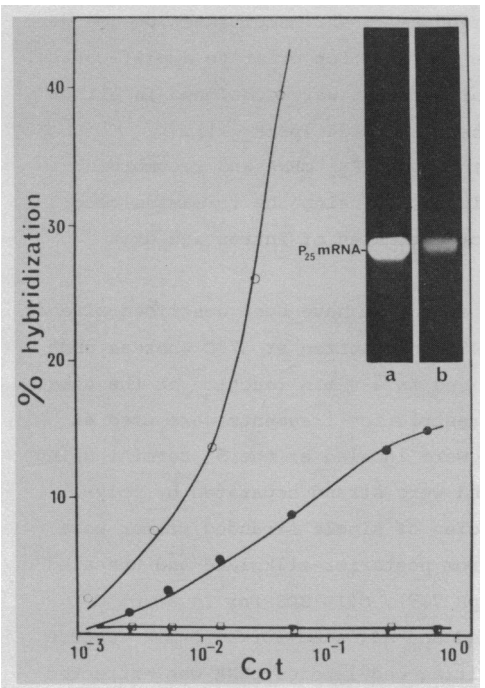

Figure 1 : Identification of BmP_{25}-1. An excess of phage DNA was hybridized with ^{32}P-labeled cDNA from non-fibroin mRNA, (●), purified fibroin mRNA (▼), and middle silkgland total mRNA (□). The Cot curve of BmP_{25}-1 alone is also shown (○). Cot is expressed as moles of nucleotide/liter/second. The inset shows blots of total PSG mRNA hybridized with cDNA from agarose gel purified P_{25} mRNA (a) and nick-translated BmP_{25}-1 DNA (b).

Eco RI digest of B. mori DNA in λCharon 4 (16). The 125,000 plaques analysed (two genome equivalents) yielded two putative positives. Further structural analysis showed that both recombinant phage shared a common genomic DNA segment, with an extra 800 bp of DNA in one of them. This suggests that the P_{25} gene, like that for fibroin (24), is present in only one copy per haploid DNA complement.

Definitive identification of the recombinants was obtained by hybridizing an excess of the phage DNAs with three kinds of ^{32}P-labeled cDNAs which were made from PSG non-fibroin mRNA, PSG fibroin mRNA, and MSG total mRNA. As shown in Fig. 1, λBmP_{25}-1 anneals at least 15% of the non-fibroin cDNA sequences but it gives no reaction with the fibroin or MSG cDNAs. In addition, the same 1100 nucleotide mRNA is found in northern blots of total PSG mRNA with a probe of either ^{32}P-labeled λBmP_{25}-1 or cDNA transcribed from agarose gel-purified P_{25} mRNA (inset in Fig. 1). Similar results were obtained with λBmP_{25}-2, thereby demonstrating that both phage recombinants contain at least part of the P_{25} gene. That λBmP_{25}-1 anneals to only 15% rather than 25-30% of non-fibroin cDNA sequences is explained by the fact that both clones are missing a portion of the 3' end of the gene (see below).

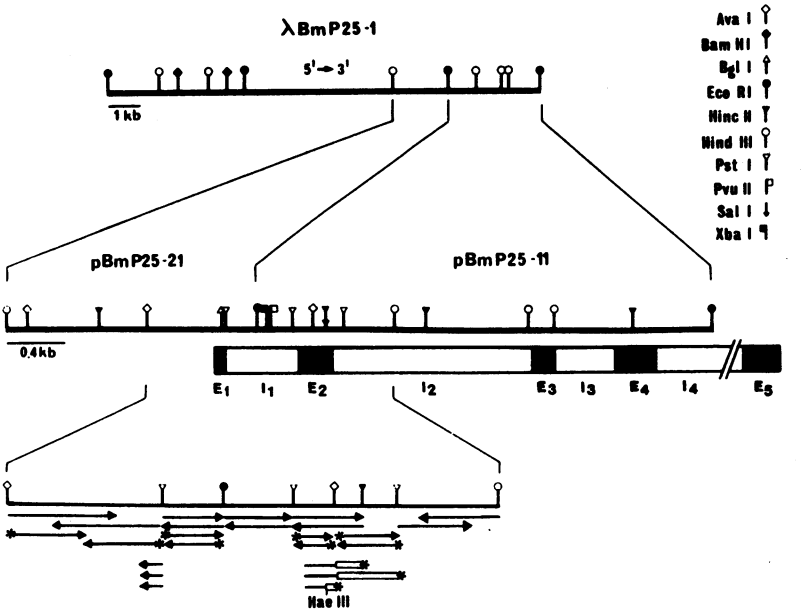

Figure 2 : Physical map of DNA from BmP_{25}-1 and related subclones $pBmP_{25}$-21 and $pBmP_{25}$-11. Exons and introns are shown as black and empty boxes respectively. The organization of the 3' region of the gene, which escaped cloning, is putative. Sequencing strategy is summarized on the bottom of the figure. Simple arrows and arrows with a star represent sequences obtained with Sanger et al (21) and with Maxam and Gilbert (20) methods respectively. Empty boxes symbolize single stranded primers used for cDNA elongation and sequencing of exon 1 and exon 2 (interrupted arrows).

Organization of the P_{25} gene

The overall organization of the P_{25} gene was deduced from structural analysis of the subcloned genomic fragments 1.6 kb Hind III-Eco RI ($pBmP_{25}$-21) and 3.1 kb Eco RI-Eco RI ($pBmP_{25}$-11). Fig. 2 shows the physical map of the DNA inserts, obtained after digestions with the enzymes listed. Annealing of blotted genomic fragments with both short and full size PSG total cDNAs (4 and 45 min of reverse transcription, respectively) allowed us to delimit the restriction fragments carrying mature mRNA sequences and to determine the polarity of the gene (data not shown). The existence of intervening sequences, which was suggested by these preliminary experiments, was further demonstrated by visualization of heteroduplexes obtained by hybridizing PSG total mRNA with DNA of the subclone $pBmP_{25}$-11. As shown in Fig. 3, the hybrids displayed two intron loops, with average lengths of 1257 $\pm$ 39 and 376 $\pm$ 23 nucleotides from the 5' to 3' ends of the mRNA. The total length of P_{25} mRNA was estimated to

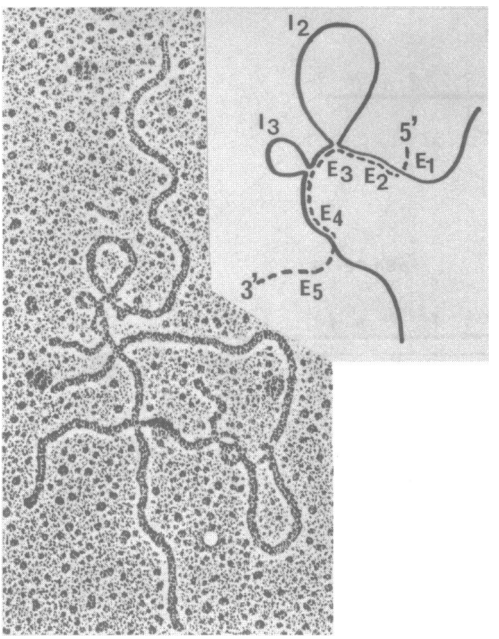

Figure 3 : Heteroduplexes between pBmP25-11 DNA and P25 mRNA. The electron micrograph shows two overlapping R-loops obtained with DNA of the clone linearized at the Bam HI site of the plasmid vector pACYC184. The inset shows the line drawing interpretation. Note the presence of 2 introns and the 5' and 3' tailing ends of the mRNA.

be 1035 ± 54 nucleotides; this agrees well with the size of 1100 nucleotides deduced from its electrophoretic mobility (1). The non-annealed stretch of P_{25} mRNA on the 5' side corresponds to exon 1 carried by $pBmP_{25}$-21 (see below); on the other side, the longer tail represents the 3' end of the mRNA, which escaped cloning, and the poly-A sequence. The size of exons 1,2 3 and 4 were estimated to be 80 ± 10, 212 ± 16, 148 ± 23 and 241 ± 18 nucleotides, respectively. Given the fact that the non-annealed portion of the mRNA on the 3' side is 405 ± 49 nucleotides, and that the mean size of the poly-A tail is 150 nucleotides (P. Couble, unpublished result), the 3' region of P_{25} gene which has not been cloned would extend to 250 nucleotides.

The 5' region of the P_{25} gene

The boundaries between exon 1 and intron 1, and between intron 1 and exon 2 were determined at the nucleotide level by comparing the nucleotide sequence of relevant genomic fragments with that of the cDNA complementary to the 5' region of P_{25} mRNA. For this purpose, different fragments were used to initiate reverse transcription close to the 5' end (see Fig. 2). For example, the 81 nucleotide Ava I-Hinc II fragment of $pBmP_{25}$-11 was isolated, end-labeled and the template strand was recovered; the cDNA resulting from the initiation with this primer was sequenced (Fig. 4). It should be mentioned that no difference

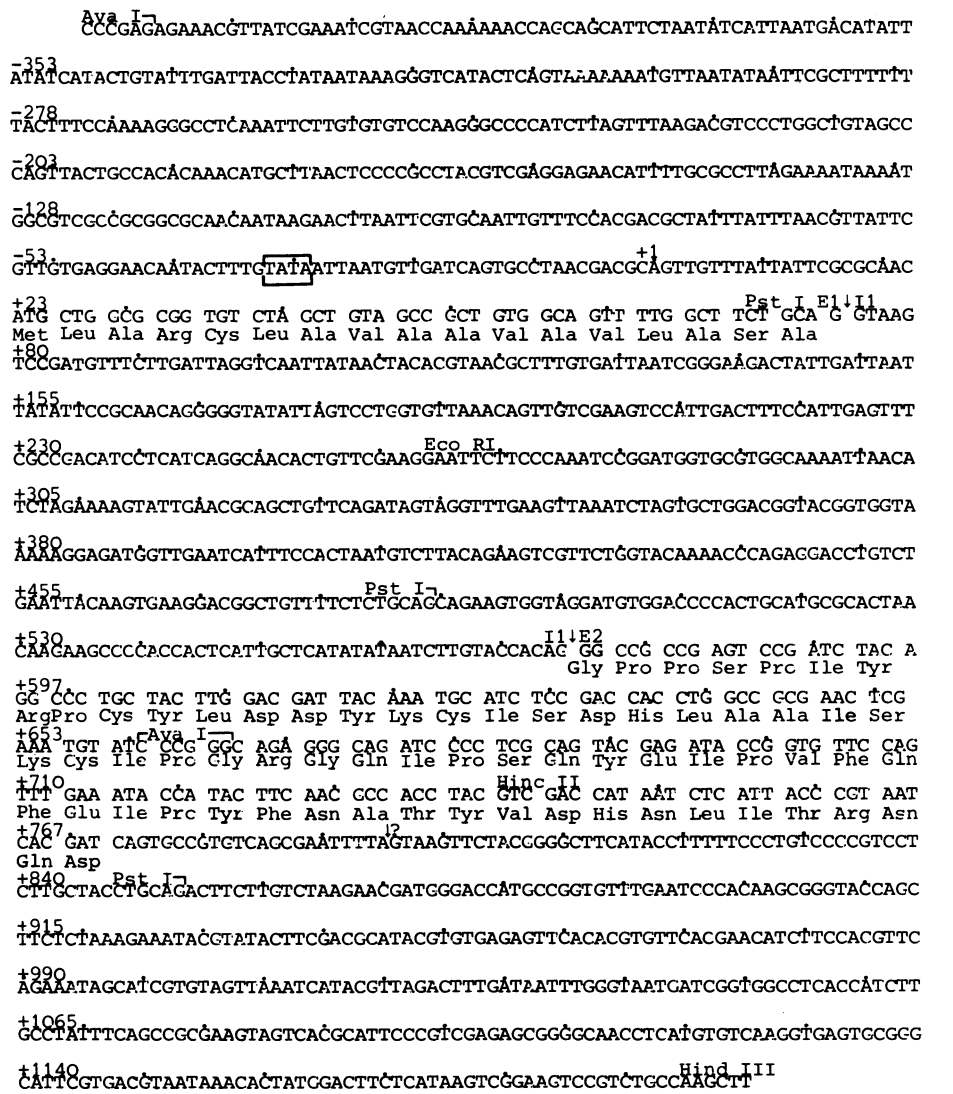

Figure 4 : Nucleotide sequence of the 5' end and of the flanking region of the P25 gene. Included in this sequence is the translation of exon 1 and of exon 2. With no definitive demonstration to ascertain the boundary between exon 2 and intron 2, the aminoacid sequence has been interrupted before reaching the first GT located at around the estimated end of the exon. The presumed boundary, located by an arrow, fulfils the consensus sequence for a donor site in intron 2. Restriction sites of relevant enzymes are indicated.

has been observed so far between the sequence of P_{25} cDNA derived from the strains 200 and 300 and that of genomic exon sequences from strain 103.

As shown in Fig. 4, the sequence of donor and acceptor sites of intron 1 GAG$^{\downarrow}$GTAAGT and CCCACAG$^{\downarrow}$GG are identical to the reported consensus sequences required for intron excision $\begin{matrix}C\\A\end{matrix}AG^{\downarrow}GT\begin{matrix}A\\G\end{matrix}$AGT and PyPyPyXCAG$^{\downarrow}G\begin{matrix}T\\C\end{matrix}$, respectively (25). The boundary between exon 2 and intron 2 has not been directly determined. The presumed boundary shown in Fig. 4, TTA$^{\downarrow}$GTAAGT, would satisfy an open reading frame over the estimated length of exon 2, and fit the consensus sequence on the intron side.

The cap site was assigned by two independent methods : by sequencing the 5' end of P_{25} mRNA as described above, and by size determination of the DNA fragments protected by P_{25} mRNA from S1 nuclease digestion. In the former experiment, cDNA made from a 24 nucleotide Hae III-Ava I primer (Fig. 2) extended to 164 and 165 nucleotides (data not shown), thereby leaving an ambiguity of a C or A residue for the mRNA start site. For the second approach, an 800 bp Hinc II-Pst I fragment of pBm P_{25}-21 was cloned in the phage M13mp10. This template was used to make a labeled copy of the DNA strand complementary to P_{25} mRNA. After annealing the labeled DNA with an excess of PSG total mRNA, the hybrids were treated with S1 nuclease. The size of the most abundant protected fragment measured on a sequencing gel was 76 nucleotides (data not shown). This locates the cap site at the A residue deduced from cDNA sequencing.

The first ATG triplet to be read in exon 1 is presumed to correspond to the initiation codon since it is the only one allowing an open reading frame over the predicted length of exon 2 (Fig. 4). The additional arguments strengthen this presumption. First, the sequence CAAC preceding this triplet agrees with the consensus sequence $\begin{matrix}C\\A\end{matrix}AA\begin{matrix}C\\A\end{matrix}$ that we deduce from the analysis of the 19 insect genes contained in the compilation of M. Kozak (26). The list includes the B.mori genes encoding fibroin and sericin which possess a CAAG and a CAAC sequence respectively (13,23). Second, the aminoacid sequence encoded by exon 1 fits the ubiquitous features of signal peptides (27,28) : a charged amino acid near the NH_2 terminus (an arginine at position 4) and a hydrophobic sequence (13 amino acids, 9 of which are valine or alanine) which are thought to initiate the export of the nascent chain across the rough endoplasmic reticulum. Whether exon 1 represents part or all the P_{25} signal peptide is unknown since the NH_2 terminus of the mature protein has not yet been determined. It is interesting to note, however, that this region of P_{25} differs from the presumed signal peptide of fibroin (13). In addition, the protein part encoded by exon 2 of P_{25} is characterized by a relatively high

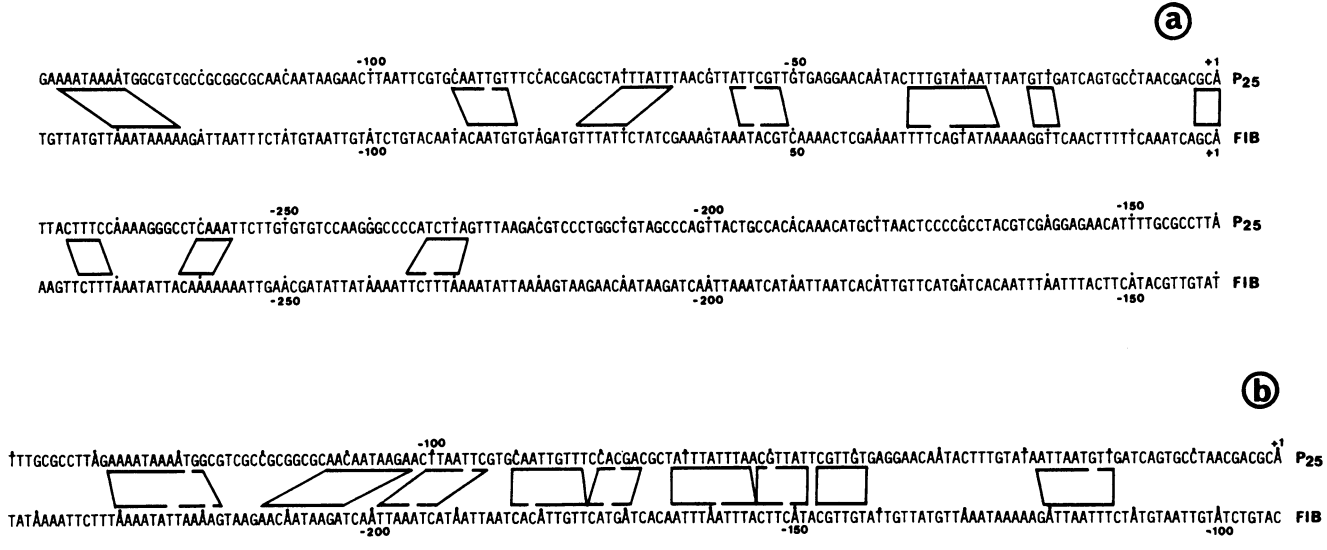

Figure 5 : Comparison of the 5' flanking sequences of the P_{25} and the fibroin genes
(a) : homologies detected in similar positions relative to the cap sites of the genes; (b) : sequences with strong homologies found at different nucleotide positions in the P_{25} and the fibroin genes.
Homologies are shown by the lines. Mismatches or deletion/insertions are indicated by gaps in the lines.

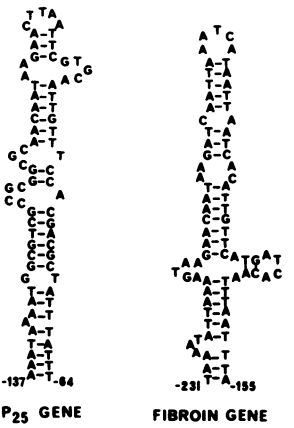

Figure 6 : Putative DNA foldings in the 5' flanking regions of the P_{25} and the fibroin genes.
Note that the stem and loop structures are included in the homologous sequences described in Fig. 5 (b).

content of proline (12%), often adjacent to isoleucine, and by a 12% content of aspartic acid and asparagine (Fig. 4).

The 5' flanking sequence of the P_{25} gene

In examining the 5' flanking sequence of the P_{25} gene, we have mainly focussed on homologies it presents with the DNA of the fibroin gene. Thus the sequence extending from -422 to the cap site of the P_{25} gene (Fig. 4) was compared to the -550 to +1 region of the fibroin gene (13). We attempted first to identify any homologies occurring in roughly similar positions relative to the cap sites of the two genes. As shown on Fig. 5(a), we detected a few blocks of homologous sequences in the regions between positions -140 and +1. They include a trinucleotide at and ahead of the cap sites. This is followed by several blocks of homology with sequences that match the canonical TATA sequence, found in most type II genes (25), and also short motifs (TTT and GTT) in the vicinity of the TATA boxes. Homologous sequences are also found further upstream, one of which is characterized by the sequence CAAT, and another by the octanucleotide AAATAAAA. Except for scarce and short sequences, we failed to detect homologous sequences in similar positions in the far upstream regions of the P_{25} and the fibroin genes.

We then looked for the existence of homologous region regardless of their position in the flanking sequences. As shown on fig. 5(b), a sequence exhibiting strong homologies was found between positions -138 and -49 of the P_{25} gene and between -230 and -142 of the fibroin gene. We found 58 matched nucleotides out of these 90 and 89 nucleotide stretches of the upstream regions of these two genes. Interestingly, a large portion of each sequence is dyad symmetric and has the potential to form a stem and loop structure. Fig. 6 shows the most stable structures that we deduced from computer

analysis. 50 nucleotides can be base-paired out of the 74 nucleotides between -137 and -64 of the P_{25} gene. In the fibroin gene, 52 nucleotides out of 76 can be base-paired between -231 and -155. We did not find any other remarkable stem and loop possibilities elsewhere in the flanking regions of the P_{25} and the fibroin genes. The possible functional significance of these sequences with regard to the regulation of expression of these genes is discussed below.

DISCUSSION

We have cloned a large portion of the gene encoding the low molecular weight silk protein of Bombyx mori and have determined its structural organization. This has permitted, in particular, a comparative sequence analysis of its 5' flanking DNA with that of fibroin gene, encoding the large silk component. Such a study is motivated by the fact that the P_{25} and fibroin genes share clear-cut stage and cell specificities, high transcription rates and also strict adjustment of transcriptional events since their mRNA are maintained in a 1:1 molar ratio during the period of silk production (1-6). The linkage of P_{25} to fibroin by disulfide bonds and the stoichiometry of their mRNAs argue for an equimolar relationship between both proteins (1).

The organization and structure of the P_{25} transcription unit in terms of size, plurality of introns, nucleotide sequence of the 5' region, do not correspond to those of the fibroin gene. This, together with the localization of the corresponding loci on two different chromosomes (29), emphasize the absence of any obvious evolutionary relationship between the genes encoding the two subunits of the major silk protein. Contrasting with this observation, it is all the more remarkable that the flanking regions of the P_{25} and fibroin genes share a high degree of sequence homology. That these may be significant with regard to the functioning of the P_{25} gene is strongly suggested by the fact that they correspond to crucial regulatory regions that have been identified in the DNA upstream of the fibroin gene. From analyses of the sequences required for in vitro transcription, Tsuda and Suzuki (14,15) have assigned a key role to two regions flanking the fibroin gene. One matches the quasi-universal promotor of RNA polymerase II transcribed gene, the sequence TATA (25), that we find in the expected position in front of the P_{25} gene. It is noticeable that not only the TATA sequence but also several motifs surrounding this promotor are shared by both genes. The other identified sequence , mainly between -239 and -116, function as a control element with strong modulator activity. Indeed, it corresponds to the region that we find highly homologous with DNA upstream of the P_{25} gene. Yet, the region of greatest homology, from -230 to -142 in the fibroin DNA, is shifted by about a hundred

nucleotides downstream in the P_{25} flanking sequences and is found between -138 and -49 (Fig. 5b). Given the observations of Tsuda and Suzuki (15) that the effect of the distal element of the fibroin gene is not abolished either by a deletion of the DNA between -65 and -44 or by an inversion of the region from -234 to -66, the shift of the equivalent sequence in P_{25} DNA may be of no consequence for its presumed role in transcription modulation. Indeed, this would be compatible with the properties of cis-acting viral enhancers (30,31) and certain eucaryotic regulatory elements (32,33) which can act at various distances from the cap site.

A common feature of the P_{25} -137 to -64 region and of the fibroin -231 to -155 sequence is that they contain dyad symmetry. Their potential stem and loop structures would be centred at nucleotides -99 and -194, respectively. This stresses the possibility that the conformational characteristics of the DNA may be decisive for the functional sequence to be recognized by regulatory proteins, so as to influence the efficiency of the "basic promotor" (the TATA box region). The two genes under study would thus be comparable to the Drosophila genes encoding heat-shock proteins, which share a conserved upstream promotor element that contains an inverted repeat (11). Dyad symmetry has also been found in the 5' flanking region of the Drosophila gene encoding the glue protein Sgs4 and which is thought to be involved in the control of transcription efficiency (34).

In a number of cases, the concerted induction of functionally related genes correlates with the presence of common short repeats in the 5' flanking regions of these genes (35,36,37). Direct evidence have been obtained by transformation experiments, on the role of such sequences in the coordinate induction of the Yeast His genes (38,39) and the Drosophila Heat Shock P70 genes (11). Consequently, sequences shared by the P_{25} and fibroin genes may be implicated in the concerted induction of their expression, which is probably hormonally mediated (1,6,40). Functional in vitro and in vivo tests are currently being developed in our laboratory to assay the properties of the above sequences on transcription modulation of the P_{25} gene. In addition, the question of whether some of these homologous sequences are unique to the P_{25} and fibroin genes is being examined by sequence analyses of other genes expressed in silkgland cells and in other tissues, such as the genes encoding sericin or cytoplasmic and muscular actins.

ACKNOWLEDGMENTS

We are indebted to Tom Eickbush who provided us with the Bombyx mori gene bank. Special thanks are due to Christian Gautier and Manolo Gouy, from

the Institut d'Evolution Moléculaire, who introduced us to computer analysis of sequences . We are also grateful to Annie Garel, Nicole Mounier and Jean-Jacques Michaille for criticism and advice throughout this work, and to Stuart Clarkson for editorial help. Part of this work was performed at the Centre de Microscopie Electronique Appliquée à la Biologie et à la Géologie of University Lyon I. These investigations were supported by the Centre National de la Recherche Scientifique (ATP No. 8176).

REFERENCES

1. COUBLE, P., MOINE, A., GAREL, A. and PRUDHOMME, J.C. (1983) Develop. Biol. 97, 398-407
2. SHIMURA, K., KIKUCHI, A., KATAGATA, Y. and OHOTOMO, K. (1982) J. Sericult. Sci. Japan 51, 20-26
3. COUBLE, P., GAREL, A. and PRUDHOMME, J.C. (1981) Develop. Biol. 82, 139-149
4. SUZUKI, Y. (1977) in Biochemical Differentiation in Insect Glands, Beerman W. Ed., Springer Verlag, New York, pp 1-44
5. SHIMURA, K., KIKUCHI, A., OHOTOMO, K., KATAGATA, Y. and HYODO, A. (1976) J. Biochem 80, 693-702
6. MAEKAWA, H. and SUZUKI, Y. (1980) Develop. Biol. 78, 394-406
7. GROSSCHEDL, R. and BIRNSTIEL, M.L. (1980) Proc. Natl. Acad. Sci. USA 77, 7102-7106
8. BENOIST, C. and CHAMBON, P. (1981) Nature (London) 290, 304-310
9. GROSVELD, G.C., ROSENTHAL, A. and FLAVELL, R.A. (1982) Nucl. Acids Res. 10, 4951-4971
10. MELLON, P., PARKER, V., GLUZMAN, Y. and MANIATIS, T. (1981) Cell 27, 279-288
11. PELHAM, H.B.R. (1982) Cell 30, 517-528
12. TSUJIMOTO, Y. and SUZUKI, Y. (1979a) Cell 16, 425-436
13. TSUJIMOTO, Y. and SUZUKI, Y. (1979b) Cell 18, 591-600
14. TSUDA, M. and SUZUKI, Y. (1981) Cell 27; 175-182
15. TSUDA, M. and SUZUKI, Y. (1983) Proc. Natl. Acad. Sci. USA 80, 7442-7446
16. EICKBUSH, T.H. and KAFATOS, F.C. (1982) Cell 29, 633-643
17. BENTON, W.D. and DAVIS, R.W. (1977) Science 196, 180-182
18. WELLAUER, P.K. and DAWID, R.W. (1977) Cell 10, 193-212
19. DAWID, R.W., SIMON, M. and DAVIDSON, N. (1971) Methods in Enzymology 21, 413-428
20. MAXAM , A.M. and GILBERT, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564
21. SANGER, F., NICKLEN, S. and COULSON, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467
22. PAPANICOLAOU, C. GOUY, M. and NINIO, J. (1984) Nucl. Acids Res. 12, 31-44
23. OKAMOTO, H., ISHIKAWA, E. and SUZUKI, Y. (1982) J. Biol. Chem. 257, 15192-15199
24. GAGE, L.P., and MANNING, R.F. (1976) J. Mol. Biol. 101, 327-348
25. BREATHNACH, R. and CHAMBON, P. (1981) Ann. Rev. Biochem. 50, 349-383
26. KOZAK, M. (1984) Nucl. Acids Res. 12, 857-872
27. BLOBEL, G. (1977) In Cell Biology, Brinkley B.R. and Porter K.R. Eds, Rockefeller University Press, pp 318-325
28. SABATINI, D.D., KREIBICH, G., MOROMOTO, T. and ADESNIK, M. (1982) J. Cell Biol. 92, 1-22
29. DOIRA, H. (1983) Sericologia 23, 245-269
30. KHOURY, G. and GRUSS, P. (1983) Cell 33, 313-314
31. WASYLYK, B., WASYLYK, C., AUGEREAU, P. and CHAMBON, P. (1983) Cell 32, 503-514

32. OLSEY, M.A. and HEREFORD, L. (1982) Proc. Natl. Acad. Sci. USA 79, 7689-7693
33. GILLIES, S.D., MORRISON, S.I., OI, V.T. and TONEGAWA, S. (1983) Cell 33, 717-728
34. MUSKAVITCH, M.A.T. and HOGNESS, D.S. (1982) Cell 29, 1041-1051
35. DAVIDSON, E.H., JACOBS, H.T. and BRITTEN, R.J. (1983) Nature 301, 468-470
36. GREZ, M., LAND, H., GIESECKE, K. and SCHUTZ, G. (1981) Cell 25, 743-752
37. DEAN, D.C., KNOLL, D.J., RISER, M.E. and O'MALLEY, B.W. (1983) Nature 305, 551-554
38. STRUHL, K. (1982) Nature 300, 284-287
39. DONAHUE, T.F., DAVIS, R.S., LUCCINI, G. and FINK, G.R. (1983) Cell 32, 89-98
40. GAREL, J.P. (1983) Experientia 39, 461-466.